

# 生物系の数理科学 第14回

山口 諒

## 機械学習における最適化と生物進化：適応度地形と勾配降下法

### 1. はじめに

機械学習の学習過程は、しばしば**最適化問題**として定式化される。一方、生物の進化もまた、**遺伝的変異**と**自然選択**が組み合わさり、適応度を高める方向へ遺伝子プールが変化していく**探索過程**と見なせる。ここでは、この2つの視点がどのように類似し、それぞれの示唆がどのように重なるかを「**適応度地形 (fitness landscape)**」と「**勾配降下法 (gradient descent)**」という概念を軸に整理する。

### 2. 適応度地形 (fitness landscape) とは

#### 2.1 進化生物学からの発想

生物進化において、ある生物集団が高い**適応度**（生存繁殖における有利性）を獲得するように遺伝的変異が蓄積していくイメージを、**地形 (landscape) の上を登る過程**にたとえるのが**適応度地形**の考え方である。生物個体の遺伝子配列や形質を「**パラメータ空間**」とみなし、その「高さ」を**適応度**として表現する。すると、

- 頂上（ピーク）は高い適応度（良い形質）を示す
- 谷や斜面は低い適応度を示す

と視覚化できる。ただし、実際には遺伝子数や形質数が莫大で次元が極めて大きくなるため、特定の組み合わせを切り取って図にすることが多い。

#### 2.2 集団遺伝学との関連

進化の過程は、「突然変異」や「組換え」で生じるさまざまな遺伝子型が、**自然選択**を介して**適応度の高い方向**へ集団の遺伝子頻度を変化させる。これは地形の斜面を“上る”とも言えるが、時には谷や平坦な尾根構造を経由するなど、複雑な経路を辿ることもある。

### 3. 勾配降下法 (gradient descent) の考え方

#### 3.1 機械学習での最適化

機械学習の典型的なタスクでは、「誤差関数  $L(\theta)$ 」や「損失関数」(loss function) を定義し、それを最小化するパラメータを求めることが目標となる。たとえば、ニューラルネットワークにおける重み・バイアスをパラメータとし、訓練データに対する誤差が小さくなるような設定が行われる。このとき、多次元のパラメータ空間において、

$$L(\theta)$$

が最小になる点を探すのが**学習**の本質である。

### 3.2 勾配降下法の手順

勾配降下法は、最適化の代表的なアルゴリズムであり、以下のステップで進む。

1. ランダムに初期パラメータ  $\theta^{(0)}$  を設定する
2. 現在の  $\theta^{(t)}$  に対して、損失関数の**勾配**  $\nabla L(\theta^{(t)})$  を計算する
3.  $\theta^{(t+1)} = \theta^{(t)} - \eta \nabla L(\theta^{(t)})$  のように、勾配方向にパラメータを更新する ( $\eta$  は学習率)
4. 反復することで、局所的・大域的な最小値に近づく

ここでは“**負の勾配方向に移動**”することで損失関数を下げるイメージが、先ほどの「適応度を上げるイメージ」と同じになる（ただし適応度を上げる場合はプラスの勾配方向に登る解釈が多いので、**最小化と最大化**で向きが逆になる点に留意）。

## 4. 類似性と相違点：進化と勾配降下

### 4.1 共通点：地形を“登る”か“降りる”か

- 生物進化：適応度地形の高いところへ向かう（最大化問題）
- 勾配降下法：損失関数（エラー）を最小化する（ $\Longleftrightarrow$  精度の向上）

いずれも「多次元空間内で最適（最大 or 最小）を探す」プロセスである。勾配を使った探索や、試行錯誤で“良さ”を改善していく点で、本質的な類似がある。

### 4.2 メタファーとしての“登山”

適応度地形や損失関数の“地形”は、多くの局所的なピークや谷が存在しうる。生物が局所ピーク（適応度の局所最大）にはまり込む場合があるように、機械学習も局所最小に捕まってしまう可能性がある。また、生物学的に中立的な変異（適応度ほぼ変わらない）に対応する平坦領域は、機械学習でいう「勾配がほぼ0の領域（サドルポイント）」にもたとえられる。

### 4.3 相違点：更新ルールと並列性

- **進化**: 個体群レベルの並列探索（多様な個体が同時に様々な遺伝的方向を試みる）。突然変異や遺伝的浮動により、実際には**勾配情報なし**でも“探索”が進む面がある。
- **勾配降下**: 通常はパラメータ1セットを更新しながら、局所微分（勾配）を指針に移動する。データ次第ではミニバッチ勾配降下など**確率的手法**も用いるが、依然として「損失関数の勾配」という局所情報を用いて進むのが中心。

進化アルゴリズム（遺伝的アルゴリズム）などは、機械学習の一部分野で「進化をまねた並列探索的な最適化手法」として使われている。